

## 航拍图像车辆检测中的圆形滤波器 HOG 特征快速计算\*

苏昂,张跃强,杨夏,于起峰

(国防科技大学 航天科学与工程学院,湖南长沙 410073)

**摘要:**航拍图像中车辆一般近似为矩形结构,因此通过统计检测窗口中的梯度方向直方图并根据梯度主方向估计车辆朝向,将检测窗口旋转到相应方向进行分类器判别。车辆检测采用级联 boosting 分类器和梯度方向直方图特征,针对旋转窗口中梯度方向直方图特征的计算,设计一种基于圆形滤波器的梯度方向直方图特征。与传统基于积分直方图的梯度方向直方图特征提取方法相比,显著提高了旋转窗口中梯度方向直方图特征的计算效率,在计算每个像素的梯度时采用查找表代替梯度向量的求模和角度计算也减小了计算量。使用真实图像进行的实验表明,该车辆检测算法快速高效。

**关键词:**车辆检测;梯度方向直方图;圆形滤波器;朝向估计;级联 boosting;航拍图像

**中图分类号:**TP391.4 **文献标志码:**A **文章编号:**1001-2486(2017)01-137-05

## Fast circle filter HOG for car detection from aerial images

SU Ang, ZHANG Yueqiang, YANG Xia, YU Qifeng

(College of Aerospace Science and Engineering, National University of Defense Technology, Changsha 410073, China)

**Abstract:** In general, cars are rectangular shape in the aerial images, so the histograms of orient gradient over the whole sliding window were computed to find the primary gradient direction and to estimate the orientation of the car in the window, and the detection window was rotated according to the car's orientation to perform classification. A cascaded boosting classifier and the HOG (histograms of orient gradient) features in the proposed car detection method were employed. To efficiently compute the HOG features in the rotated window, a fast HOG features extraction method based on CFHOG (circle filter based histograms of orient gradient), which was more efficient than the classical HOG extraction method based on integral histograms. In addition, look-up tables are used to speed up the calculation of the orientation partition and magnitude. A set of experiments on real images prove the applicability and high efficiency of the proposed car detection method.

**Key words:** car detection; histograms of orient gradient; circle filter; orientation estimation; cascaded boosting; aerial images

航拍图像中的车辆检测在城市规划和交通管理等方面有着重要的应用需求,一直以来都是研究的热点,然而高效、高鲁棒性的车辆检测依然是一个富有挑战性的问题。

特征和分类器是车辆检测算法的两个重要内容,常用特征有 Haar-like 特征<sup>[1-2]</sup>、梯度方向直方图特征<sup>[3]</sup> (Histograms of Orient Gradient, HOG)、局部二值模式特征<sup>[4]</sup> (Local Binary Pattern, LBP) 和尺度不变特征变换<sup>[5]</sup> (Scale Invariant Feature Transform, SIFT) 等,常用分类器主要有 boosting 和支持向量机 (Support Vector Machine, SVM) 等。Grabner 同时采用 Haar-like, HOG 和 LBP 三种特征,并结合一种在线 boosting 分类器进行车辆检测<sup>[6]</sup>。Kembhavi 采用了包括 HOG 特征在内的多种特征并使用部分最小二乘

(Partial Least Squares, PLS) 算法对原始高维特征进行降维,然后使用 SVM 分类器进行车辆检测<sup>[7]</sup>。Tuermer 采用 HOG 特征并基于 Adaboost 算法训练车辆检测分类器,在预处理步骤中通过分割出车辆的潜在区域以提高检测效率同时降低虚警率<sup>[8]</sup>。Moranduzzo 通过估计待检测图像上提取的 HOG 特征与车辆样本的 HOG 特征之间的相似性来检测车辆<sup>[9]</sup>。HOG 特征因其良好的性能成为应用最为广泛的特征之一,而 boosting 分类器相对 SVM 而言更为简单且计算速度更快,因此,本文基于 HOG 特征和级联 boosting 分类器进行车辆检测。

车辆在图像中可能有各种朝向,一种方法是针对不同朝向的车辆分别训练多个分类器并在每个滑动位置使用这些分类器分别进行检测<sup>[7-8]</sup>,

\* 收稿日期:2015-08-17

基金项目:国家自然科学基金资助项目(61402489)

作者简介:苏昂(1987—),男,河南扶沟人,博士研究生,E-mail:suang2008@126.com;

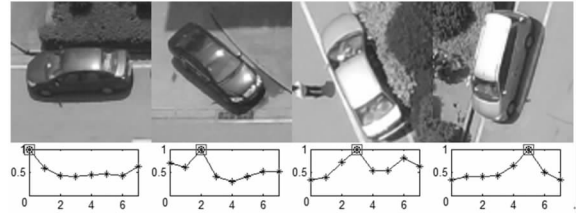
杨夏(通信作者),男,副研究员,博士,E-mail:mr\_honson@sina.com

另一种方法是针对特定朝向的车辆训练一个分类器并在每个滑动位置间隔一定角度旋转检测窗口进行检测<sup>[6]</sup>。为了避免在同一个位置进行多次检测,首先根据窗口中的梯度主方向估计车辆朝向,然后旋转检测窗口提取 HOG 特征,再进行分类器判别。传统基于积分直方图的 HOG 特征快速计算仅适用于与图像坐标轴平行的矩形块<sup>[10]</sup>。通过旋转图像或者采用多角度积分图像,可以实现多角度旋转窗口中倾斜矩形块的 HOG 特征计算,但计算过程较为复杂<sup>[11-12]</sup>。为了快速提取旋转窗口中的 HOG 特征,设计了一种基于圆形滤波器的 HOG 特征。该方法适用于任意朝向旋转窗口,且特征提取只需取值运算,无须加法运算,能够显著提高 HOG 特征提取效率。此外,采用查找表代替复杂的梯度向量求模和角度计算也提高了算法效率。使用真实图像进行的实验表明,本文方法在不损失精度的情况下能够显著提高车辆检测效率。

## 1 车辆朝向估计

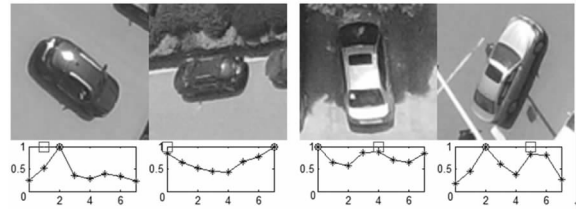
航拍图像中车辆具有各种朝向,传统的在同一位置进行多次检测的方法检测效率不高。车辆在图像上一般为近似矩形结构,在车身方向将存在一个梯度主方向。而车身方向与边缘梯度方向垂直,因此将得到的原始梯度方向直方图平移  $\pi/2$ ,使得  $0^\circ \sim 180^\circ$  对应的直方图分块编号  $0 \sim 7$  分别对应车身朝向  $0^\circ, 22.5^\circ, 45^\circ, 67.5^\circ, 90^\circ, 112.5^\circ, 135^\circ$  和  $157.5^\circ$  方向。使用滑动窗口提取特征时,统计窗口的梯度方向直方图并找到最大值,则该方向对应车身朝向。为了验证该结论,从真实图像中选取了 471 个车辆样本,将人工标注的车身朝向与整个窗口的梯度方向直方图作对比,结果如图 1 所示,其中方框表示人工标注的车身朝向,圆圈表示检测窗口梯度方向直方图的最大值。统计所有车辆样本窗口梯度方向直方图最大值与人工标注车身朝向之间的偏差,结果见表 1。如图 1(a) 所示,332 个样本直方图最大值与车身朝向重合,如图 1(a) 所示,97 个样本直方图最大值与车身朝向相差 1 个方向块,统计显示直方图最大值与车身朝向相差不超过 1 个方向块的样本合计占总样本的 90.87%。如图 1(b) 和图 1(c) 所示,45 个样本直方图最大值与车身朝向相差 1 个方向块以上,占总样本的 9.13%。该结果表明检测窗口中梯度方向直方图的最大值大概率对应车身方向。因此,本文进行分类器训练和车辆检测时,首先统计滑动窗口的 8 方向梯度

方向直方图并找到最大值对应的方向,然后旋转窗口至该方向提取 HOG 特征。采用该方法,不仅避免了在同一位置进行多次检测而减小了计算量,也使得标注训练样本时不必标注车辆朝向并训练多个分类器而减小了工作量。由于车辆朝向估计存在误差,用于分类器训练的样本中也包括了少量朝向估计有误差的车辆样本,使得车辆检测分类器包含了对车辆朝向估计误差的适应性。



(a) 正确的车辆朝向估计

(a) Correct car orientation estimation



(b) 估计偏差 22.5

(c) 偏差 22.5 以上

(b) Estimation bias 22.5

(c) Bias above 22.5

图 1 基于梯度方向直方图的车辆朝向估计

Fig. 1 Car orientation estimation based on histograms of orient gradient

表 1 梯度主方向与人工标注车辆朝向对比

Tab. 1 Comparison of the primary gradient direction and labeled orientation of the car

方向偏差	0	1	2	3	4
样本数量	332	96	10	25	10
百分比/%	70.49	20.38	2.12	5.10	1.91
合计/%	90.87		9.13		

## 2 HOG 特征快速计算

### 2.1 基于积分直方图的 HOG 特征快速计算

HOG 特征一般采用矩形区域块 (rectangular block) 构造描述子,每个区域块由  $2 \times 2$  的矩形小块 (rectangular cell) 组成,将  $0^\circ \sim 180^\circ$  划分为 8 块计算梯度方向直方图,最终得到的一个 32 维的 HOG 特征描述向量<sup>[3]</sup>。对于这种矩形 HOG 特征 (Rectangular Histograms of Orient Gradient, RHOG),通常采用积分直方图进行快速计算。如图 2 所示,文献<sup>[12]</sup>为计算旋转窗口中的 RHOG

特征,将原图旋转  $22.5^\circ$  得到旋转图,分别计算原图和旋转图的  $0^\circ$  积分直方图和  $45^\circ$  积分直方图,以计算  $0^\circ \sim 180^\circ$  区间 8 种朝向旋转窗口中的 RHOG 特征,具体如下所述:对于  $0^\circ$  和  $90^\circ$  旋转窗口在原图的  $0^\circ$  积分直方图上计算,对于  $45^\circ$  和  $135^\circ$  旋转窗口在原图的  $45^\circ$  积分直方图上计算,对于  $22.5^\circ$  和  $112.5^\circ$  旋转窗口在旋转图的  $0^\circ$  积分直方图上计算,对于  $67.5^\circ$  和  $157.5^\circ$  旋转窗口在旋转图的  $45^\circ$  积分直方图上计算。提取一个 32 维的 RHOG 特征,共需要  $3 \times 32$  次加法运算和  $4 \times 32$  次取值操作。在预处理阶段,需要 1 次图像旋转操作、 $2 \times 8$  次  $0^\circ$  积分图像运算和  $2 \times 8$  次  $45^\circ$  积分图像运算。

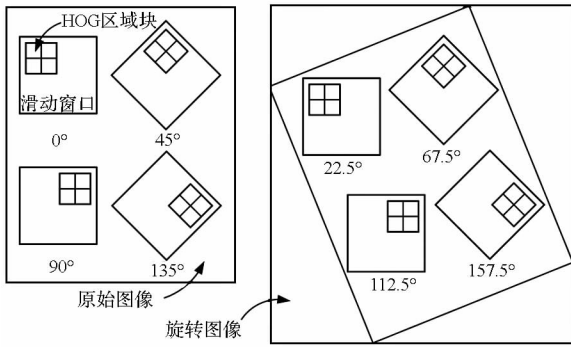


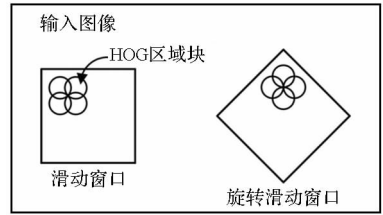
图 2 基于积分直方图的旋转窗口 RHOG 特征快速计算  
Fig. 2 RHOG features fast computation based on integral image

## 2.2 基于圆形滤波器的 HOG 特征快速计算

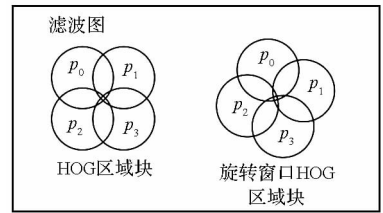
为了快速提取旋转窗口中 HOG 特征,本文提出一种基于圆形滤波器的 HOG 特征(Circle Filter Histograms of Orient Gradient, CFHOG)计算方法,该方法采用  $2 \times 2$  的圆形小块(circular cell)组成区域块,如图 3(a)所示。同样将  $0^\circ \sim 180^\circ$  划分为 8 块来统计梯度方向直方图,使用元素全为一的圆形模板构建滤波器,然后将圆形滤波器分别与输入图像的 8 个方向梯度图做卷积,得到各梯度方向的滤波图,则滤波图上每个像素即为以该点为圆心的圆形小块所对应的梯度方向直方图统计值。如图 3(b)所示,设已知检测窗口中每个 CFHOG 特征的四个圆形小块在滤波图上对应的圆心坐标分别为  $p_0 \sim p_3$ ,根据该坐标依次在各梯度方向滤波直方图上取值即可直接得到该 CFHOG 特征。

基于积分直方图的传统 RHOG 特征快速计算方法的本质是利用积分图像计算矩形区域内各方向梯度的累加和,而基于圆形滤波器的 CFHOG 特征快速计算方法的本质是利用圆形滤波器计算

圆形区域内各方向梯度的累加和。本文方法提取一个 32 维的 CFHOG 特征,仅需要 32 次取值操作,在预处理阶段,需要 8 次滤波运算。与基于积分直方图的传统 RHOG 特征快速计算方法相比具有以下优势:①预处理和特征提取运算更加简洁,特征提取仅需取值操作,无须加法运算,提高了特征提取的效率;②能够快速计算任意角度旋转窗口中的 HOG 特征,而传统基于积分直方图的方法对旋转窗口的适应能力有限。



(a) CFHOG 特征设计  
(a) CFHOG feature design



(b) CFHOG 特征快速提取

(b) Fast extraction of CFHOG feature

图 3 基于圆形滤波器的 CFHOG 特征

Fig. 3 CFHOG features based on circle filter

## 2.3 基于查找表的梯度快速计算

为了统计 HOG 特征,在每个像素位置通常使用一维模板  $[-1 \ 0 \ 1]$  分别计算图像  $x, y$  方向的梯度,然后合成梯度向量并计算该向量的模和角度。为加速 HOG 特征提取,可利用查找表代替复杂的求模运算和反正切运算<sup>[13]</sup>。对于灰度图像,  $x, y$  方向梯度的值域区间为  $[-255 \ 255]$ ,因此构建 4 个  $511 \times 511$  的查找表,前两个表中存储梯度向量投影的相邻两个方向块编号,后两个表中存储梯度向量在相邻两个方向块中的投影分量。计算出  $x, y$  方向的梯度后,分别作为查找表的行列索引,直接取值即可得到该点对梯度方向直方图的贡献。

## 3 级联 boosting 分类器

使用文献[12]中的级联 boosting 分类器进行车辆检测,分类器表达式为:

$$H(x) = \text{sign} \left[ \sum_{i=1}^T \alpha_i h_i(x) \right] \quad (1)$$

其中,  $\mathbf{x}$  表示输入数据,  $H(\mathbf{x})$  表示强分类器,  $T$  为弱分类器数量,  $t$  为弱分类器编号,  $\alpha_t$  为弱分类器权重,  $h_t(\mathbf{x})$  表示弱分类器, 是一个简单的阈值型函数, 可表示为:

$$h_t(\mathbf{x}) = \text{sign}[D_1(\mathbf{f}_t(\mathbf{x})) - D_0(\mathbf{f}_t(\mathbf{x}))] \quad (2)$$

其中,  $\mathbf{f}_t(\mathbf{x})$  表示输入数据  $\mathbf{x}$  中编号为  $t$  的弱特征, 而  $D_1(\mathbf{f}_t(\mathbf{x}))$  和  $D_0(\mathbf{f}_t(\mathbf{x}))$  分别表示该弱特征到正聚类中心和负聚类中心的距离, 设正负聚类中心分别满足高斯分布  $N(\boldsymbol{\mu}_{1t}, \boldsymbol{\Sigma}_{1t})$  和  $N(\boldsymbol{\mu}_{0t}, \boldsymbol{\Sigma}_{0t})$ , 则距离测度表达式为:

$$D_i(\mathbf{f}_t(\mathbf{x})) = \ln|\boldsymbol{\Sigma}_{it}| + (\mathbf{f}_t(\mathbf{x}) - \boldsymbol{\mu}_{it})^T \boldsymbol{\Sigma}_{it}^{-1} (\mathbf{f}_t(\mathbf{x}) - \boldsymbol{\mu}_{it}) \quad (3)$$

其中:  $i = 0, 1$ ;  $|\cdot|$  表示矩阵的行列式。

为了构建级联结构, 设  $c_i(\mathbf{x}) = \alpha_i h_i(\mathbf{x})$  表示第  $i$  个弱分类器的输出值,  $H_t(\mathbf{x})$  表示前  $t$  个弱分类器输出的累加分数, 可表示为:

$$H_t(\mathbf{x}) = \sum_{i=1}^t c_i(\mathbf{x}) \quad (4)$$

对所有弱分类器训练完成后, 选择分类错误率最低的若干个组成最终的级联 boosting 分类器, 其中每个弱分类器为一层。然后通过样本训练给每层设定一个拒绝阈值  $r_t$ , 使得当  $H_t(\mathbf{x}) < r_t$  时, 直接判别输入样本为非目标不再进入下一层的判别; 当  $H_t(\mathbf{x}) \geq r_t$  时, 继续进行后一层的弱分类器判别。由于图像中大部分位置不是目标, 这种级联分类器的早期拒绝机制能够大大加速检测算法的效率。

## 4 实验

### 4.1 图像数据集与实验配置

在城市中四处楼顶位置使用相机拍摄地面车辆得到 26 幅图像数据。图像分辨率为  $1000 \times 562$  像素, 图像中车辆长度从 26 ~ 54 像素不等。车辆检测窗口设为  $42 \times 42$  像素, 以 1.25 为尺度缩放系数构建三层图像金字塔进行分层检测以适应车辆尺度变化, 则每层的等效检测窗口大小依次为  $42 \times 42$ ,  $52 \times 52$  和  $66 \times 66$  像素。检测窗口中 HOG 特征的小块尺寸设为  $6 \times 6$  像素, 每个区域块包括  $2 \times 2$  个小块, 滑动步长为 3 像素, 那么在尺寸为  $42 \times 42$  像素的检测窗口中共包含 121 个 HOG 特征, 设每个 HOG 特征对应一个弱分类器, 最终选择分类错误率最低的 60 个弱分类器组成级联 boosting 分类器。

利用数据集 26 幅图像上标注的车辆正样本以及随机选取的负样本, 采用 5 折交叉验证测试

本文车辆检测算法, 算法使用 C++ 编程实现, 实验计算机使用 Windows 8.1 64 位操作系统, 配置 8 G 内存以及 Intel Core i7 8 × 2.6 Hz CPU。

### 4.2 车辆检测结果

本文算法的部分车辆检测结果如图 4 所示, 图中方框表示检测结果, 圆圈表示测试集上标注的车辆样本真值。统计本文基于 CFHOG 的车辆检测算法的精度和召回率曲线<sup>[14]</sup>并与文献[12]基于 RHOG 的车辆检测算法作对比, 如图 5 所示。本文基于 CFHOG 的车辆检测算法与文献[12]基于 RHOG 的算法每帧耗时对比结果见表 2。

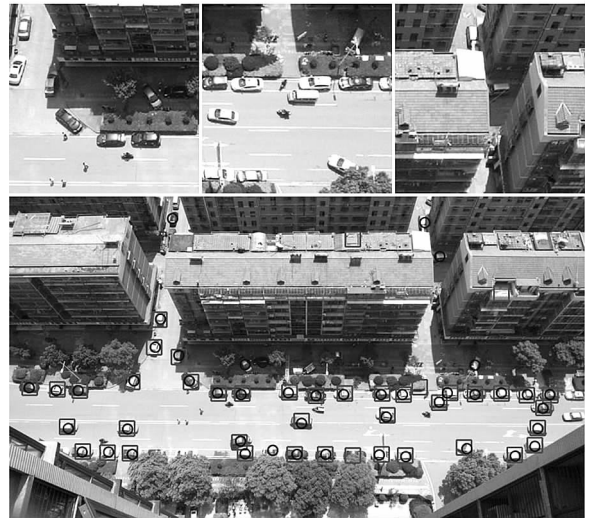


图 4 车辆检测结果示例

Fig. 4 An example of the car detection results

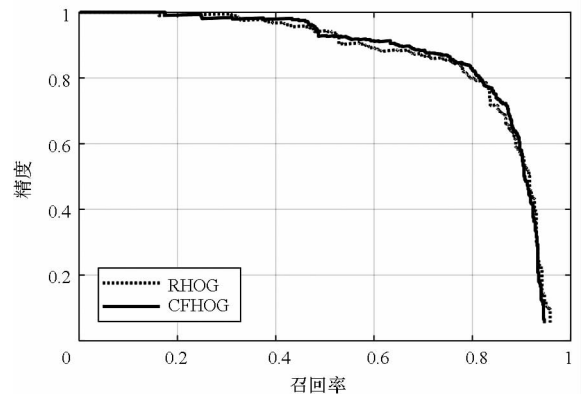


图 5 车辆检测算法精度和召回率曲线

Fig. 5 Precision-recall curves of the car detection algorithms

结果表明本文算法与文献[12]算法检测效果相当, 而在检测速度方面占优。在图像预处理阶段本文算法比传统基于 RHOG 的检测算法快约 0.192 s, 而在分类器判别阶段, 由于本文提出的 CFHOG 特征提取只需取值操作, 使得算法耗

时比传统基于 RHOG 的检测算法少约 0.860 s。最终本文基于 CFHOG 的车辆检测算法比文献[12]传统基于 RHOG 的检测算法每帧快 1.052 s。

表 2 车辆检测算法每帧耗时

Tab. 2 Average computational time of the airplane detection algorithm for each frame

算法	RHOG	CFHOG
预处理/s	0.593	0.401
检测/s	2.817	1.957
合计/s	3.410	2.358

## 5 结论

为了快速计算旋转窗口中的 HOG 特征,本文提出一种基于圆形滤波器的 HOG 特征快速提取方法,该方法使用圆形小块代替传统方法的矩形小块,并在图像预处理阶段利用圆形滤波器统计各圆形小块的梯度方向直方图,与文献[12]中传统基于积分直方图的 HOG 特征提取方法相比:①预处理和特征提取运算更加简洁,特征提取仅需取值操作,无须加法运算,提高了特征提取的效率;②能够快速计算任意角度旋转窗口中的 HOG 特征。最后,使用与文献[12]相同的分类器和图像测试集进行的实验表明,采用 CFHOG 特征的车辆检测算法在不损失检测精度的条件下显著提高了车辆检测算法的效率。

## 参考文献 (References)

[1] Viola P, Jones M J. Robust real-time face detection [J]. *International Journal of Computer Vision*, 2004, 57 (2): 137 - 154.

[2] Lienhart R, Maydt J. An extended set of haar-like features for rapid object detection [C]//*Proceedings of International Conference on Image Processing, ICIP*, 2002.

[3] Dalal N, Triggs B. Histograms of oriented gradients for human

detection [C]//*Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR*, 2005.

[4] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(7): 971 - 987.

[5] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, 60(2): 91 - 110.

[6] Grabner H, Nguyen T T, Gruber B, et al. On-line boosting-based car detection from aerial images [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2008, 63 (3): 382 - 396.

[7] Kembhavi A, Harwood D, Davis L S. Vehicle detection using partial least squares [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33 (6): 1250 - 1265.

[8] Tuermer S, Kurz F, Reinartz P, et al. Airborne vehicle detection in dense urban areas using HOG features and disparity maps [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2013, 6(6): 2327 - 2337.

[9] Moranduzzo T, Melgani F. Detecting cars in UAV images with a catalog-based approach [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52 (10): 6356 - 6367.

[10] Porikli F. Integral histogram: a fast way to extract histograms in cartesian spaces [C]//*Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005(1): 829 - 836.

[11] Messom C H, Barczak A L. Fast and efficient rotated haar-like features using rotated integral images [C]//*Proceedings of Australian Conference on Robotics and Automation*, 2006.

[12] Su A, Sun X L, Liu H B, et al. Online cascaded boosting with histogram of orient gradient features for car detection from unmanned aerial vehicle images [J]. *Journal of Applied Remote Sensing*, 2015, 1(9): 096063.

[13] Yan J J, Lei Z, Wen L Y, et al. The fastest deformable part model for object detection [C]//*Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2014.

[14] Agarwal S, Awan A, Roth D. Learning to detect objects in images via a sparse, part-based representation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(11): 1475 - 1490.